LOSS MODELS

ACTS 457

Prof. Akoto Yaw Omari-Sasu

KNUST

January 15, 2026

Course Overview

This course introduces students to advanced Loss Models, emphasizing the application of frequency and severity models in real-world business contexts, particularly in the insurance industry. Students are required to understand how to quantify and model losses and apply these models to address business problems.

Course Objectives

After this course, the students are expected to;

1. Differentiate between the various types of deductibles (ordinary and franchise) and analyze how each type impacts loss payments in different insurance scenarios.
2. Estimate the effects of external factors such as inflation rate, coinsurance factors, and coverage limits on insurance policies.
3. Analyze the aggregate loss and aggregate payments for a portfolio of individual risks, considering different risk profiles, loss distributions, and policy structures.

COURSE OUTLINE

- Review of Probability Distributions
- Unit 1: Claim Severity Models
- Unit 2: Exposure and Coverage Modification
- Unit 3: Empirical Distributions
- Unit 4: Aggregate Payments Models
- Unit 5: Discrete Distributions

Books

- Loss Models from data to decisions, Wiley Series
- Models for Quantifying Risk, ACTEX Academic Series
- Construction and Evaluation of Actuarial Models, 7th edition by Abraham Weishaus

Probability Distributions

Fundamental Continuous Distributions

- The Exponential Distribution
  The density function of the exponential distribution in addition to its moments are given below;

$$f(x) = \frac{1}{\theta} e^{-\frac{x}{\theta}}, \quad \text{for} \quad x \geq 0, \quad \theta > 0$$

$$\text{Mean} = \theta, \quad \text{Variance} = \theta^2$$

- The Normal Distribution

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad -\infty < x < +\infty$$

$$\text{Mean} = \mu, \quad \text{Variance} = \sigma^2$$

- The Pareto Distributions
  The Pareto distribution can be classified in two forms as a double-parameter Pareto and a single-parameter Pareto. The Pareto distributions are defined for all positive values of x by the density function.

Double-Parameter Pareto

$$f(x) = \frac{\alpha\theta^{\alpha}}{(x+\theta)^{\alpha+1}} \quad \alpha > 0, \ \ \theta > 0$$

$$F(x) = 1 - \left(\frac{\theta}{x+\theta}\right)^{\alpha}$$

$$\text{Mean} = \frac{\theta}{\alpha - 1}, \quad \text{Variance} = \frac{\alpha\theta^2}{(\alpha-2)(\alpha-1)^2} \quad \alpha > 1, \quad \alpha > 2$$

Single-Parameter Pareto

$$f(x) = \frac{\alpha\theta^{\alpha}}{x^{\alpha+1}} \quad \alpha > 0, \ \ \theta > 0$$

$$F(x) = 1 - \left(\frac{\theta}{x}\right)^{\alpha}$$

the kth moment is given by: $E(X^k) = \frac{\alpha\theta^k}{\alpha - k} \quad \alpha > 1$

Fundamental Discrete Distributions

- The Poisson Distribution

$$f(x) = \frac{\lambda e^{-\lambda}}{x!} \quad \lambda > 0, \ \ x = 0, 1, 2, 3...$$

$$\text{Mean} = \lambda \quad \text{Variance} = \lambda$$

- The Binomial Distribution

$$P(N = k) = \binom{m}{k} p^k (1-p)^{m-k} \quad \text{for} \quad k = 0, 1, 2, \ldots$$

$$E[N] = mp \quad \text{and} \quad Var[N] = mp(1-p)$$

Claim Severity Models

Modifications of The Loss Random Variable

In this section, we assume that a single loss event is represented by the severity random variable X, and we define the random variable Y to represent the amount paid (claim amount) to offset part or all of that loss by the insurance company. When an individual is partly responsible for a loss(i.e., when $Y<X$), they are less likely to engage in risky behavior, knowing they will bear some of the consequences if something goes wrong. Warranties and insurance policies strictly make Y less than X.

Deductible: A deductible is a threshold amount, denoted d, which must be exceeded by a loss in order for a claim to be paid. There are two main categories of deductibles, namely; Ordinary deductibles and Franchise deductibles.

Ordinary Deductibles: An Ordinary deductible is a threshold amount, denoted d, which must be exceeded by a loss in order for a claim to be paid. Thus the amount paid when the deductible is exceeded (X >d) is Y = X – d. Hence the amount of claim random variable Y is defined as

$$Y = (X - d)_+ = \begin{cases} 0 & \text{if } X \leq d \\ X - d & \text{if } X > d \end{cases}$$

where the notation $(X - d)_+$ is read "the excess of X over d, if positive". From the equation above, it is clearly seen that when $X \geq d$, Y = 0 and thus we get the discrete version of Y as

$$P_y(0) = Pr(Y \geq 0) = F_x(d)$$

The continuous version of the distribution of Y for y>0 is also given by

$$f_y(y) = f_x(d + y)$$

The moments of the random variable Y are given

$$E(Y^k) = E((X-d)_+)^k = \int_d^\infty (x-d)^k f_x(x)\,dx$$

if X is continuous,

$$E(Y^k) = E\left((X-d)_+\right)^k = \sum_{x \geq d} (x-d)^k p_x(x)$$

if X is discrete,

The random variable for the amount not paid due to the deductible is the minimum of X and d .This random variable is denoted $X \wedge d$. Its expectation is referred to as **the limited expected value**. Given this, the expected **payment per loss** for an ordinary deductible can be simplified to get

$$E(Y) = E(X) - E(X \wedge d)$$

where E(X) is the expected value of the loss event

$$E\left[(X \wedge d)^k\right] = \int_0^d x^k f_X(x) + d^k[1 - F(d)]\, dx$$

$$E\left[(X \wedge d)^k\right] = \int_0^d k x^{k-1} S_X(x)\, dx = \int_0^d S_X(x)\, dx$$

The amount Paid per payment event (Mean excess loss), for which we need to define a third ramdom variable Z. Whenever there is a payment, (that is Y $>$ 0) then Z is thesame as Y. Thus this can be written as Z = Y |Y $>$ 0 which implies Z = Y |X $>$d. NB:(Y = X-d)

$$Z = \begin{cases} Underfined & \text{if } X \leq d \\ X - d & \text{if } X > d \end{cases}$$

The kth moment of Z if X is continuous is given by

$$E[Z^k] = E[Y^k | Y > 0] = E[(X - d)^k | X > d]$$

$$E[Z^k] = \frac{\int_d^\infty (x - d)^k f_x(x)\, dx}{1 - F_x(d)}$$

and the discrete is given by

$$E[Z^k] = \frac{\sum_{x \geq d}(x-d)^k p_X(x)}{1 - F_x(d)}$$

Therefore the kth moment for the payment per payment event can be deduced from the equations above as

$$E[Z^k] = \frac{E[Y^K]}{1 - F_x(d)}$$

The expected value of Z can thus be expressed as

$$E[Z^k] = \frac{[E(X) - E(X \wedge d)]}{1 - F_x(d)}$$

The payment per loss can also be expressed in terms of the payment per payment event as

$$E[Y] = E[Z] \times [1 - F_x(d)]$$

Example One

The severity distribution function of claims data for automobile property damage coverage SIC is given by an exponential distribution

$$F(x) = 1 - e^{\frac{-x}{5000}}$$

To improve the likelihood of this policy, a deductible of 500 per claim for the next period was imposed by SIC. Calculate E(Y) and E(Z) for the next period

Answer

$$S(x) = e^{\frac{-x}{5000}}$$

The expected payment per loss is

$$E(Y) = \int_d^{\infty} S_x(x)\, dx = \int_{500}^{\infty} e^{\frac{-x}{5000}}\, dx$$

$$E(Y) = \int_d^{\infty} S_x(x)\, dx = \int_{500}^{\infty} e^{\frac{-x}{5000}}\, dx$$

$$E(Y) = 5000 \times e^{-\frac{500}{5000}} = 4524.097$$

Also the expected payment per payment is given by

$$E[Z] = \frac{E[Y]}{1 - F_x(d)} = \frac{E[Y]}{1 - F_x(500)}$$

$$E[Z] = \frac{4524.097}{1 - e^{-\frac{500}{5000}}} = 4999.9$$

Example Two

From example one, calculate the uncovered loss, E(X $\wedge d$).

Answer

$$E(X \wedge d) = E(X) - E(Y)$$

$$E(X \wedge d) = 5000 - 4524.097 = 475.903$$

Try Question One

Calculate the expected payment per loss for insurance coverage with deductible of 5 if the loss distribution follows a single-parameter Pareto with parameter $\alpha = 2$ and $\theta = 1$

Try Question Two

Losses have the following distribution;

$$F(x) = 1 - 0.4e^{-\frac{x}{20}} - 0.6e^{-\frac{x}{2000}}$$

Insurance coverage is subject to a deductible of 100. Determine the average payment per payment and the average payment per loss.

Try Question Three

Which of the deductibles will you prefer should you be running an insurance company; ordinary of franchise and why?

Franchise Deductible: A franchise deductible, d, means that if the loss is below d, nothing is paid, but if the loss is d or higher, the full amount is paid.

The expected value of the payment per loss is then the same as for an ordinary deductible, plus an additional payment of d times the probability of exceeding the deductible which is 1- F(d) or in other words S(d).

The expected value of the payment per payment for a franchise deductible is the expected payment per payment for an ordinary deductible plus d. In summary, the franchise deductible modifies the ordinary deductible by adding the deductible when there is a positive amount paid.

Thus the payment per loss variable is given by

$$Y = \begin{cases} 0 & \text{if } X \leq d \\ X & \text{if } X > d \end{cases}$$

The payment per payment variable is also given by;

$$Z = \begin{cases} Undefined & \text{if } X \leq d \\ X & \text{if } X > d \end{cases}$$

From the definitions, The expected payment per loss is defined;

Example One

For a franchise deductible d find the expected payment per loss in each of the following cases

- X is exponential with mean θ
- X is uniform on (0,θ), where d $< \theta$

Answer

- If d is a franchise deductible, then the expected payment per loss is

$$E[Y^*] = E[(X-d)_+] + d[1 - F_x(d)] = \theta e^{-\frac{d}{\theta}} + de^{-\frac{d}{\theta}} = (\theta + d)e^{-\frac{d}{\theta}}$$

- If d is a franchise deductible, then the expected payment per loss is

$$E[Y^*] = \frac{(\theta - d)^2}{2\theta} + d[1 - \frac{d}{\theta}] = \frac{\theta^2 - d^2}{2\theta}$$

Try Question One

Losses have the following distribution;

$$F(x) = 1 - 0.3e^{-\frac{x}{20}} - 0.7e^{-\frac{x}{2000}}$$

Insurance coverage is subject to a franchise deductible of 150. Determine the average payment per payment and the average payment per loss.

Exposure and Coverage Modification To The Loss Random Variable

The Loss Elimination Ratio and Inflation

The Loss Elimination Ratio: It is the ratio of the decrease in the expected payment with an ordinary deductible to the expected payment without the deductible. It is the proportion of the loss which the insurer doesn't pay as a result of the deductible. While many types of coverage modifications can decrease the expected payment, the term loss elimination ratio is reserved for the effect of changing the deductible. The Loss Elimination Ratio is given by the expression below;

$$LER(d) = \frac{E(X) - E(Y)}{E(X)} = \frac{E(E \wedge d)}{E(X)}$$

Provided the expected value of the loss exists.

Example One

The amount of a single claim has a Pareto distribution with parameter $\alpha = 3$ and $\theta = 2000$. Calculate the Loss Elimination Ratio for a deductible of 500

Answer

This is a double parameter Pareto distribution, thus the formula is given by;

$$E[X] = \frac{\theta}{\alpha - 1}$$

$$E(X \wedge d) = \frac{\theta}{\alpha - 1}\left[1 - \left(\frac{\theta}{d + \theta}\right)^{\alpha - 1}\right]$$

$$E(X \wedge d) = 1 - \left(\frac{\theta}{d + \theta}\right)^{\alpha - 1} = 1 - \left(\frac{2000}{2500}\right)^{2} = 0.36$$

The Effects of Inflation

When inflation is taken into consideration, it decreases the loss elimination ratio. Inflation increases costs, but it turns out that when there is a deductible the effect of inflation is magnified. Thus if Y = (1 + r)X, where X is the original variable, Y is the inflated variable and r is the inflation rate, then the expected value can be calculated as follows

$$E[Y \wedge d] = E[(1+r)X \wedge d] = E[X \wedge \frac{d}{1+r}]$$

Therefore the formula for the expected payment per loss is given by

$$E[Y] = (1+r)[E(X) - E(X \wedge \frac{d}{1+r})]$$

Also the expected payment amount per payment is given by;

$$E[Z] = \frac{(1+r)[E(X) - E(X \wedge \frac{d}{1+r})]}{1 - F_x(\frac{d}{1+r})}$$

Example One

The underlying loss distribution function for a certain line of business in 1991 is;

$$F(x) = 1 - x^{-5}$$

From 1991 to 1992, 10% inflation impacts all claims uniformly.Determine the 1992 LER for a deductible of 1.2.

Answer

We could see that the distribution follows a single parameter Pareto after examining it.

$$LER(d) = \frac{E[X \wedge \frac{d}{1+r}]}{E[X]}$$

$$E[X] = \frac{\alpha\theta}{\alpha - 1} = \frac{5}{5-1} = \frac{5}{4} = 1.25$$

$$E\left[X \wedge \frac{d}{1+r}\right] = \frac{\alpha\theta}{\alpha-1} - \frac{\theta^{\alpha}}{(\alpha-1)\left(\frac{d}{1+r}\right)^{\alpha-1}}$$

$$E\left[X \wedge \frac{1.2}{1.1}\right] = \frac{5}{4} - \frac{1^5}{(4)\left(\frac{1.2}{1.1}\right)^{5-1}} = 1.07348$$

$$LER = \frac{1.07348}{1.25} = 0.858784$$

Try Question

You are given the following;

- Losses follow a Pareto distribution with parameters $\theta = n$ and $\alpha = 2$, where n is a constant. There is deductible of 2n.

1. What is the Loss Elimination Ratio
2. Over a period of time, inflation has uniformly affected all losses, causing them to double, but the deductible remains the same. Calculate the new LER

Coverage Modifications

Policy Limit: The policy limit is defined as the most that would be paid by an insurance company. The typical policy limit arises in a contract where for losses below u, the insurance pays the full loss but for loss greater than u, the insurance pays only u. i.e.,

$$Y = (X \wedge u) = \begin{cases} X & \text{if } X \leq u \\ u & \text{if } X > u \end{cases}$$

In the absence of a deductible there will be a payment associated with every loss event, so the random variable for amount of payment per payment event, denoted Z, will be the same as the random variable for the amount of payment per loss event, which is Y. i.e Z= Y= (X$\wedge u$)

$$E[Y] = u - \int_0^u F_X(x)\,dx = \int_0^u [1 - F_X(x)]\,dx$$

$$E\left[(X \wedge u)^k\right] = \int_0^u x^k f_X(x) + \int_u^\infty u^k f_X(x)\,dx$$

For the effect of inflation **r** on policy limit, the expected value of the loss is

$$(1+r)E[X \wedge \frac{u}{1+r}]$$

Example One

A limit of 3000 has been imposed on a Pareto distribution with $\alpha = 3$ and scale parameter $\theta = 2000$.Determine the expected payment per loss with the limit and also when a uniform inflation of 10% is applied.

Answer

Before inflation,For the Pareto distribution, the expected loss is

$$E[X \wedge 3000] = \frac{2000}{2}[1 - (\frac{2000}{2000+3000})^2] = 480$$

After Inflation

$$1.1E[X \wedge \frac{3000}{1.1}] = (1.1)\frac{2000}{2}[1 - (\frac{2000}{2000+\frac{3000}{1.1}})^2] = 903.11$$

Coinsurance Factor

The final common coverage modification is coinsurance. In this case the insurance company pays a proportion, α, of the loss and the policyholder pays the remaining fraction. Coinsurance of α means that a portion, α, of each loss is reimbursed by insurance. If coinsurance is the only modification, (in the absence of deductible) this changes the loss variable X to the payment variable; $Z = Y = \alpha X$

For a policy with limit u, deductible d, and coinsurance factor α,

$$E[Y] = \alpha[E(X \wedge U) - E(X \wedge d)]$$

When inflation is added to the above three factors, we have the payment per loss as

$$E[Y] = \alpha(1+r)[E(X \wedge \frac{u}{1+r}) - E(X \wedge \frac{d}{1+r})]$$

For the payment per payment expected value, we have

$$E[Z] = \frac{E[Y]}{1 - F_x(\frac{d}{1+r})}$$

Try Question

Losses in 1998 follow a two parameter Pareto distribution with parameters $\alpha = 0.5$ and $\theta = 1000$. Insurance pays claims up to a maximum covered loss of 20,000 subject to a deductible of 500 with 75% coinsurance. Determine the size average claim payment.

Empirical Distributions And Bonuses

Empirical Distributions

- For well-established lines of business offering ample credible loss data, directly using an empirical distribution may serve more effectively than fitting a parametric model to those data. The random variable X is treated as discrete with the observed loss amounts treated as equally likely. This makes the calculation of the PDF, CDF, and moments of X relatively simple to do.
- Furthermore, for some risk products such as warranties, insurance, or default protections, we can expect the availability of data sets involving many thousands of observations. In situations as this, empirical distributions can be both more effective and much simpler to use than parametric models.

Example

A manufacturer of washing machines provides a repair warranty on each new machine it sells. Under the warranty, all repairs are paid by the company except for a deductible of \$50 per repair that is paid by the customer. The historical total repair costs are shown in the following table.

Event	Loss Amount(x)
1	150
2	85
3	25
4	200
5	72
6	182
7	52
8	332
9	141
10	281
11	171
12	253
13	33
14	197
15	273
16	96

Determine each of the following:

1. The probability function and cumulative distribution function of X
2. The expected payment amount made by the company per loss event and the expected payment amount per payment event.
3. The variance in the amount paid by the company per loss event
4. The impact of 20% uniform inflation on the expected payment amount per loss event.

Bonus Payments

Suppose that X is a loss random variable. Suppose that a risk manager receives a bonus equal to the amount by which the loss is less than some specified limit u; if the loss is above u, there is no bonus. We can thus formulate the bonus a function for the bonus as

$$Bonus = \begin{cases} u - X & \text{if } 0 < X \leq u \\ 0 & \text{if } X > u \end{cases}$$

To solve these problems, express the bonus in terms of the earned premiums; normally it always takes the form max [0, c(rP-X)], where r is the loss ratio, P is earned premium, X is losses, and c is some constant. We can then factor out crP and write

$$crP - cmin(rP, X) = crP - c(X \wedge rP)$$

Aggregate Payments Models

Individual Risk Versus Collective Risk

There are two standard approaches for developing a model for the sum S of all losses or claims during a particular period. In the first approach, for Contract i, where $i = 1, 2, \ldots, n$, a severity distribution is chosen to model the loss amount, denoted by X_i, for that contract during the period. We assume

- The X_i's are independent, but not necessarily identically distributed random variables. Different insureds could have different distributions of aggregate losses. Since an insured may not submit any claims, we have $\Pr(X_i = 0) > 0$.
- n is the size of the group and is a fixed number.

This model is called the individual risk model and is given by the equation $S = \sum_{i=1}^{n} X_i$ with first two moments given as

$$E[S] = \sum_{i=1}^{n} E[X_i]$$

$$Var[S] = \sum_{i=1}^{n} var[X_i]$$

Where claim amounts are identically distributed under the various contracts, the moments of S becomes

$$E[S] = nE[X]$$

$$Var[S] = nVar[X]$$

Where the mean, E[X] and Variance, Var[X] are identical given each observation.

The second approach is to keep track of losses by claim incurred, rather than by policy. Let X_i for $i = 1, 2, \ldots, N$, be the severity random variables, the sizes of the N losses. If S is the aggregate loss variable, then the aggregate claim amount for the entire portfolio is given by

$$S = \sum_{i=1}^{N} X_i \quad \text{and this approach is called } \textbf{the collective risk model}$$

The following **assumptions** were made.

1. The X_i's are independent identically distributed random variables. That is to say, every claim size has the same probability distribution and is independent of any other claim size.
2. The X_i's are independent of N. The claim counts are independent of the claim sizes. The first two moments of the collective risk model are written as:

$$E[S] = E[N]E[X]$$

$$Var[S] = E[N]Var[X] + (E[X])^2 Var[N]$$

For a compound poisson distribution, the variance formula is reduced to

$$Var[S] = \lambda Var[X] + \lambda E[X]$$

$$Var[S] = \lambda E[X^2]$$

Example

You are given the following:

- An insured's claim severity distribution is described by an exponential distribution $F(x) = 1 - e^{-\frac{x}{1000}}$.
- The insured's number of claims is described by a negative binomial distribution with parameters $r = 2$ and $\beta = 2$.
- A \$500 per loss deductible is in effect.

Calculate the standard deviation of the aggregate losses in excess of the deductible.

Answer

The frequency of losses has a negative binomial distribution with $r = 2$ and $\beta = 2$. However, only losses above 500 are paid. The probability that a loss will be paid is $e^{-\frac{d}{1000}} = e^{-\frac{500}{1000}} = e^{-0.5}$. The distribution of the number of paid claims is then negative binomial with parameters $\beta = 2e^{-0.5}$ $r = 2$. The distribution of individual losses above the deductible is exponential with mean 1000, since the exponential distribution is memoryless. Let N be the distribution of paid claims, and X be the distribution of individual losses above the deductible. Then

$$E[X] = 1000$$

$$Var[X] = 1000$$

$$E[N] = 2 \times 2e^{-0.5} = 2.42612$$

$$Var[N] = (2 \times 2e^{-0.5})(1 + 2e^{-0.5}) = 5.36916$$

$$Var[S] = 5.36916(1000)^2 + 2.42612(1000) = 7,795,281$$

Therefore the standard deviation of S is

$$\sqrt{7,795,281} = 2792$$

Try Question

An insurance company sold 100 fire insurance policies. The probability of claim per policy is 0.05. Given that the claim amount for each policy is U[0,400] and the probability of more than one claim is 0. Calculate the mean and variance of the aggregate claim.

Try Question

Assume there is a chance of 0.2 that there is a claim. When the claim occurs, the loss is exponentially distributed with $\lambda = 0.5$

1. Find the mean and variance of the claim distribution.
2. Suppose there are 500 independent policies with this loss distribution. Compute the mean and variance of their aggregate claim.

Discrete Distributions

The purpose of this unit is to introduce a large class of counting distributions. Counting distributions are discrete distributions which have probabilities on only the non - negative integers (0, 1, 2, 3 . . .). In the context of insurance, counting distributions describe the number of events of such losses to the insured or claims to the insurance company. We shall therefore be looking at random variables that are commonly used to model annual frequency of losses in a collective risk model for aggregate annual losses. The most commonly employed frequency models are Poisson, negative binomial and binomial

Frequency Models

The annual frequency of losses for a portfolio of risks and the corresponding annual frequency of claims payments are examples of counting distributions. The probability function for a counting distribution is generally denoted in one of two ways:

$$f_N(n) = Pr(N = n)$$

The three distributions mentioned earlier, are the only counting distributions that satisfy the (a, b, 0) class property. There are constants a and b such that

$$Pr(N = n) = (a + \frac{b}{n})Pr(N = n - 1) \quad \text{for} \quad n = 1, 2, 3..$$

Using the property (a,b,0) repeatedly, we can recursively generate the values of $Pr(N = 1)$, $Pr(N = 2)$, if the starting value $Pr(N = 0)$ is known.

The Poisson Distribution

$$f(x) = \frac{\lambda e^{-\lambda}}{k!} \quad \lambda > 0, \quad k = 0, 1, 2, ...$$

The Negative Binomial Distribution

$$P(N = k) = \binom{k + r - 1}{k} \left(\frac{1}{1+\beta}\right)^r \left(\frac{\beta}{1+\beta}\right)^k \quad \text{for} \quad k = 0, 1, 2, \ldots$$

$$E[N] = r\beta, \quad Var[N] = r\beta(1+\beta)$$

The Binomial Distribution

$$P(N = k) = \binom{m}{k} p^k (1-p)^{m-k} \quad \text{for} \quad k = 0, 1, 2, \ldots$$

$$E[N] = mp \quad \text{and} \quad Var[N] = mp(1-p)$$

(a,b,0) Class of Distribution

Let P_k be the probability function of a discrete random variable. It is a member of the (a, b, 0) class of distributions provided that there exists constants a and b such that

$$\frac{P_k}{P_{k-1}} = a + \frac{b}{k} \quad , k = 1, 2, 3$$

The table below shows the starting value for the recursion P_0

Distribution	a	b	P_0
Poisson	0	λ	$e^{-\lambda}$
Binomial	$\frac{-p}{1-p}$	$\frac{(m+1)p}{1-p}$	$(1-p)^m$
Negative Binomial	$\frac{\beta}{1+\beta}$	$\frac{(r-1)\beta}{1+\beta}$	$(1+\beta)^{-r}$
Geometric	$\frac{\beta}{1+\beta}$	0	$(1+\beta)^{-1}$

Example

The claim frequency follows a distribution in the $(a, b, 0)$ class. You are given the following probabilities:

- Probability of 4 claims: $P(4) = 0.066116$
- Probability of 5 claims: $P(5) = 0.068761$
- Probability of 6 claims: $P(6) = 0.068761$

Calculate the probability of no claims.

Answer

First solve for a and b,

$$\frac{P(5)}{P(4)} = a + \frac{b}{5} = \frac{0.068761}{0.066116}$$

$$\frac{P(6)}{P(5)} = a + \frac{b}{6} = 1$$

Solving the two equations simultaneously, a $= 0.8$ and b $= 1.2$. Since both a and b are positive, the distribution follows the negative binomial.

$$\frac{\beta}{1+\beta} = 0.8 = a$$

$$\beta = 4$$

$$\frac{(r-1)\beta}{1+\beta} = 1.2 = b$$

$$r = 2.5$$

Therefore,

$$P_0 = (1+\beta)^{-r} = (1+4)^{-2.5} = 0.01789$$

Example

X is a discrete random variable with a probability function which is a member of the (a, b, 0) class of distributions. You are given:

- $P(X=0) = P(X=1) = 0.25$
- $P(X=2) = 0.1875$

Calculate $P(X=3)$

Answer

$$\frac{P_1}{P_0} = a + \frac{b}{1} = 1$$

$$\frac{P_2}{P_1} = a + \frac{b}{2} = \frac{0.1875}{0.25} = 0.75$$

$$a = 0.5$$

$$b = 0.5$$

Using the relation

$$\frac{P_3}{P_2} = a + \frac{b}{3} = \frac{P_3}{0.1875}$$

$$0.5 + \frac{0.5}{3} = \frac{P_3}{0.1875}$$

Therefore

$$P_3 = 0.125$$

Try Question

For a Frequency distribution in the (a,b,1) class.You are given that

- $P_k = 0.0768$
- $P_{k+1} = P_{k+2} = 0.08192$
- $P_{k+3} = 0.0768432$

Determine the mean and the variance for this distribution.

(a,b,1) Class of Distribution

In the previous unit, we looked at the (a,b,0) distribution. In this unit, we shall be looking at a situation where the (a,b,0) is inapplicable. The (a,b,1) is a modified version of the (a,b,0) distribution. Ways of modeling P_0 and P_1

- Zero Truncated
- Zero Modified

Zero Truncated

Assume that

$$P_0^T = 0$$

and

$$P_1^T = \frac{P_1}{1 - P_0}$$

but for k $\geq$ 2,

$$\frac{P_k^T}{P_{k-1}^T} = a + \frac{b}{k}$$

Zero Modified

Assume that

$$P_0^M = c$$

and

$$P_1^M = \frac{(1 - P_0^M)P_1}{1 - P_0}$$

$$P_1^M = (1 - P_0^M)P_1^T$$

but for k $\geq$ 2,

$$\frac{P_k^M}{P_{k-1}^M} = a + \frac{b}{k}$$

Example

For a distribution of the (a, b, 1) class, $P_1 = 0.4$, $P_2 = 0.2$ and $P_1 = 0.1$ Determine P_0

Answer

$$\frac{P_2^M}{P_1^M} = a + \frac{b}{2} = 0.5$$

$$\frac{P_3^M}{P_2^M} = a + \frac{b}{3} = 0.5$$

$$a = 0.5$$

$$b = 0$$

This is a zero-modified geometric distribution, since b = 0 implies the distribution is geometric.

$$a = 0.5 = \frac{\beta}{1+\beta}, \quad \beta = 1$$

And for geometric

$$P_0 = \frac{1}{1+\beta} = 0.5$$

$$P_1 = \frac{\beta}{(1+\beta)^2} = 0.25$$

This implies

$$P_1^M = \frac{(1 - P_0^M)P_1}{1 - P_0} = 0.4$$

$$\frac{(1 - P_0^M)0.25}{1 - 0.5} = 0.4$$

$$1 - P_0^M = 0.8$$

$$P_0^M = 0.2$$

THANK YOU